IntuitionLabs

INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

eCTD Validation Requirements: A Comprehensive Technical Guide

> Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-01-12 · ectd · ectd validation · regulatory submissions · fda · ema · regulatory affairs · common technical document · ich m2

Original article: https://intuitionlabs.ai/articles/ectd-validation-requirements

In this article

1. Executive Summary
2. Introduction and Background
3. The eCTD Specification and Structure
4. Basic Architecture and Components
5. eCTD Validation: Goals and Processes
6. Validation Tools and Industry Practice
7. Regulatory Perspectives and Requirements
8. Data and Trends in eCTD Usage and Validation
9. Case Studies and Examples
10. Discussion: Implications and Future Directions
11. Quality and Efficiency Trade-offs
12. Trends and Future Outlook
13. Conclusion

[Revised May 6, 2026] This article has been reviewed and updated to reflect the current state of eCTD validation as of mid-2026, including the ongoing global rollout of eCTD v4.0 (FDA acceptance for new applications since Sep 16, 2024; EMA optional use for new centrally authorised product MAAs since Dec 22, 2025), Health Canada Validation Rules v5.3, and EU Module 1 v3.1.1 / Validation Criteria v8.2 (mandatory since Dec 2025).

Executive Summary

The **electronic Common Technical Document (eCTD)** has become the global standard for regulatory submissions of pharmaceutical and biologic products to health authorities, replacing older paper or non-standard electronic formats. Developed under the auspices of the International Council for Harmonisation (ICH) in the early 2000s, the eCTD specifies a hierarchical file structure (organized into Modules 1–5) with an XML “backbone” index that references all constituent documents (typically in PDF). This structured approach has enabled harmonization of dossier content and expedited review workflows. However, it imposes strict **validation requirements**: each submission must conform precisely to the technical specification (folder organization, file naming, index content, data formats, linking of references, etc.) and to agency-specific business rules (e.g. correct use of Module 1 regional tags, presence of metadata, allowable file formats).

This report provides an exhaustive analysis of eCTD validation requirements: the historical context and evolution of e-submissions, the technical specification of eCTD, the variety of rule sets used by regulatory authorities (FDA, EMA, Health Canada, PMDA, etc.), the common errors and validation failures observed in practice, and the tools and processes sponsors use to ensure compliance. We review data on submission volumes and rejection rates, industry experiences, and the regulatory implications of validation rigor. For example, FDA data indicate that although only about **2%** of submissions are ultimately rejected for technical non-compliance (en.ennov.com), FDA’s gateway statistics count millions of transactions across all centers and transaction types, so that rate cannot simply be multiplied out into an eCTD rejection count (www.fda.gov). We

include illustrative scenarios of validation challenges (e.g. missing or mismatched backbone files triggering technical rejection) and note the substantial downstream costs for industry; one software vendor estimates that a delayed application can cost up to **$8 million per day** (a vendor estimate, not an FDA figure) (www.docshifter.com).

Key findings include: health authorities routinely publish and update detailed **validation rule sets** (for example, Health Canada's eCTD Validation Rules v5.3 contains hundreds of specific checks, such as file size limits and folder naming conventions (www.canada.ca) (www.canada.ca)); sponsors are strongly advised to pre-validate with commercial validation software (FDA recommends validating before filing, and Health Canada encourages commercially available tools), since FDA uses "commercial off-the-shelf" validators (e.g. Lorenz eValidator) to check each dossier (www.accessdata.fda.gov) (en.ennov.com); and submissions with high-severity validation errors are technically rejected and must be corrected and resubmitted, a step distinct from a later *Refuse to File* decision made during filing review (FDA issues notices identifying which error codes are enforced as high severity (www.fda.gov) (en.ennov.com)).

In the **future**, the eCTD format continues to evolve. FDA began accepting eCTD v4.0 for new applications in September 2024 (www.fda.gov), and EMA has accepted it on an optional basis for new centrally authorised product MAAs since December 22, 2025, with prior contact with its eCTD v4.0 team required (esubmission.ema.europa.eu) (esubmission.ema.europa.eu). New rules (e.g. ICH M8/eCTD v4) introduce additional metadata and content types (e.g. expanded controlled vocabularies). These developments imply further changes to validation engines and sponsor practices. At the same time, there are opportunities for improved interoperability (linking eCTD to real-world electronic data, integration with emerging regulatory information management systems, etc.). The report concludes with a discussion of the implications of stringent validation – including cost/benefit tradeoffs – and recommended future directions, such as leveraging automation to manage the growing complexity of e-submission rules.

Introduction and Background

In the late 20th century, as drug development data volumes exploded, regulators worldwide recognized the need for standardized dossier formats. The Common Technical Document (CTD) was agreed under ICH in 2000 (pharmafile.com), harmonizing submission structure across regions (including Modules 1–5 for administrative, summaries, quality, nonclinical, and clinical data). However, the CTD itself was format-agnostic – it defined the outline of content but not how it should be delivered. Concurrently, advances in computing and the Internet drove the vision of fully electronic submissions. In 1994 the **Multi-Agency Electronic Regulatory Submission (MERS)** project laid groundwork for electronic data exchange, and after CTD harmonization the focus shifted to "Electronic Standards for the Transfer of Regulatory Information (ESTRI)" (pharmafile.com). This effort culminated in the **electronic CTD (eCTD)** standard, developed by ICH's M2 Expert Working Group and released in final form in the mid-2000s (pharmafile.com) (www.ema.europa.eu). The eCTD specifies that dossiers be delivered as digital documents (typically PDF) organized in a pre-defined folder and XML-indexed format.

As an industry summary explains, *"The eCTD defines a structured compilation of electronic documents arranged in specific folders and referenced from a hierarchical data file in XML format"* (pharmafile.com). Each document in an eCTD is included once and referenced by path and metadata, rather than transmitted

repeatedly. The format is inherently multiregional: Module 1 of the eCTD is regional (for example, US-specific "us-regional.xml" or EU-specific "eu-regional.xml"), while Modules 2–5 follow the international ICH specification (often called ICH M4/M2 for the backbone and file specifications) (pharmafile.com) (studylib.net). Notably, **PDF** is mandated for human-readable documents, to ensure consistency; the eCTD rules generally forbid submitting Word or other proprietary formats.

By the late 2000s, major agencies had phased in eCTD. FDA began accepting eCTD submissions in the 2000s, although eCTD did not become mandatory for NDAs, ANDAs and BLAs until May 5, 2017, and for commercial INDs and master files until May 5, 2018 (www.fda.gov) (pharmafile.com). Similarly, the European Medicines Agency (EMA) made eCTD mandatory for electronic submissions in the centralised procedure from January 1, 2010 (pharmafile.com). In North America, Health Canada and the US FDA gradually phased in eCTD with PDUFA performance goals: for example, FDA policy documents anticipated that 24 months after final guidance (circa 2009 for NDAs/BLAs) all major submissions must be eCTD (www.accessdata.fda.gov).Canada issued Module 1 guidance by 2012 (www.canada.ca), and required ever more submissions in eCTD format in subsequent years. Other regulators, including Japan's PMDA, have since adopted their own eCTD requirements.

By 2025, eCTD is essentially the **only** accepted submission format for most high-level regulatory filings in the US, EU, Canada, and Japan. Other formats (such as the older **NeeS** format used in Europe or paper) are deprecated or allowed only in special cases. This ubiquity makes the burden of *validation* extremely important: regulators have instituted *rigorous automated checks* that an eCTD conforms exactly to the published technical specifications and business rules. The remainder of this report dives into those validation requirements and their implications.

The eCTD Specification and Structure

Basic Architecture and Components

An eCTD submission is a self-contained folder (or archive) with a defined top-level naming scheme (often `DocketName/0000` for the initial sequence, then `0001`, `0002`, etc. for subsequent updates). Each sequence folder contains:

- A single **XML backbone file** (e.g. `index.xml` per ICH, and a regional file such as `us-regional.xml` or `eu-regional.xml`) that defines the content hierarchy and links to all documents included in that sequence.
- All the supporting **documents** (typically PDF, but sometimes images, spreadsheets, etc., subject to allowed formats), organized into the tree of subfolders that mirror the CTD module/section structure.
- Where applicable, **Study Tagging Files (STF XML)**, which group the documents of each nonclinical or clinical study; FDA requires them for study files in specified Module 4 and Module 5 sections, and omitting one can trigger a validation error.

The backbone XML declares *leaf nodes* and *branch nodes* of the submission tree. A leaf node corresponds to a particular document (e.g. a PDF file) and is associated with attributes (title, filename, etc.) in the XML. For example, an entry in `index.xml` might look like this (simplified):

```
<leaf>
 <file id="F3" location="M5-5-4/Study_Report.pdf"/>
 <title>Study 0012 Report</title>
 <subject>...</subject>
 <revision-history ...
</leaf>
```

This ensures each document is indexed. Modifications or continuity across sequences are performed by setting the *`operation` attribute* on each leaf (new, append, replace or delete), implementing the lifecycle of the dossier. Proper use of these actions is critical: for example, trying to "replace" a file that does not exist in the previous sequence, or skipping sequence numbers (e.g. submitting 0003 without 0002), are validation errors (www.canada.ca) (www.canada.ca).

Modules 2–5 follow the global CTD content outlines (Quality, nonclinical, clinical), but Module 1 is region-specific. For instance, in US eCTDs the `us-regional.xml` includes administrative forms (like FDA Form 356h), cover letters, FDA-specific narratives (e.g., establishment names, marketing history), etc. In EU submissions, `eu-regional.xml` includes the forms for a Marketing-Authorisation Application (MAA) and other EU-specific content. Each region also has its own *schemata*, *stylesheets*, and *controlled vocabularies* to ensure consistency in the XML format.

File Formats and Naming: eCTD rules strictly specify permitted file types (for example, FDA accepts PDF versions 1.4 through 1.7, PDF/A-1 and PDF/A-2, which must be text searchable and carry no security settings or password protection (www.fda.gov)) and naming conventions (no spaces, restrictions on characters, uniqueness of filenames). Even details like capitalizing file extensions can matter in some loci. The FDA's eCTD Implementation Guide and technical conformance guide document, for example, define exact allowable MIME types and extension lists. Health Canada's v5.3 rules flag invalid extensions as errors (www.canada.ca). Leaf titles should be short, descriptive and meaningful, while the CTD section headings themselves are controlled by ICH (and by regional specifications for M1). The FDA published a *Table of Contents Headings and Hierarchy* to standardize the labels used in Module 1 folders such as `M1-3-3-1` for Specimen Labeling, etc. (It recently updated that Toc list in Feb 2025 (www.fda.gov)).

Lifecycle and Versioning: Each submission (often called a "sequence" or "transaction") builds on the last. The folder names are numeric: typically the initial submission is `0000`, then an amendment or response is `0001`, etc. Health Canada's rules require sequences to increment by 1 without gaps (www.canada.ca). Each file in a new sequence retains the same name if unchanged, or is replaced/deleted according to its action. The index also includes a unique `doc-id` GUID for each submission, which regulators use to track the dossier across sequences (and alignment of data in their systems). **Publishing tools** automatically generate these GUIs and maintain the publishing metadata. If any part of this structure is off (for example, an out-of-order sequence folder name), the validation catches it (www.canada.ca) (www.canada.ca).

In summary, the technical specification defines a very rigid structure: correct order of files, correct metadata tags, and exact formatting (PDFs must use an accepted PDF version, XML must be well-formed and valid against schemas). If any file is missing, misnamed, inaccessible, or contains unallowed content, a validation error will be raised. The next sections discuss in detail how these specifications are enforced, and what sponsors must do to comply.

eCTD Validation: Goals and Processes

Validation in the eCTD context means an electronic check of a submission package against a set of rules and specifications. The primary goal is to ensure the submission is *complete*, *self-consistent*, and *compliant* with the agreed technical standards before it enters substantive review. Validation has two facets: (1) **Structural validation** – verifying that the dossier's XML, folder hierarchy, and file properties adhere to the technical spec; and (2) **Content and business rule validation** – verifying that mandatory elements are present (e.g. required M1 forms, appendices) and that they make sense (e.g. file sizes not too large, no duplicate IDs).

Regulatory agencies perform automated validation on every eCTD. The FDA, for example, has conducted eCTD validations for well over a decade (en.ennov.com). The FDA uses a *commercial off-the-shelf* (COTS) validation tool to execute the structural checks (www.accessdata.fda.gov) (en.ennov.com). Companies are expected to use a comparable tool internally *before* submission to catch issues early. The FDA even hints at this in training materials: *"FDA uses a COTS product to validate eCTD submissions. For more information, review Specifications for eCTD Validation Criteria."* (www.accessdata.fda.gov). In other words, sponsors cannot rely on their own judgment alone; they must pass the same tests that the regulator will apply.

Typically, eCTD validation occurs in stages. **Pre-submission**: sponsors use vendor tools (from companies like Lorenz, EXTEDO, EMPOWER, PAREXEL, etc.) to run a validation against the official criteria for the target region. These tools produce reports listing errors and warnings. Industry consultants emphasize that because agencies update validation rules frequently (e.g. see the many version notes on FDA's site (www.fda.gov)), it is *very important* to use the *latest* criteria and software version before filing (www.onixls.com).

Upon submission, the electronic gateway of each agency automatically receives the package and validates it. If *high-severity* errors are found, the dossier is **technically rejected** and must be corrected and resubmitted. This technical rejection is separate from a *Refuse to File* (RTF) decision, which FDA makes later, during filing review. The FDA, for instance, classifies errors as High, Medium, or Low severity (en.ennov.com). A High-severity error means the eCTD is invalid at a basic level (e.g. missing core files) and the submission is considered not received. The agency's policy (as reiterated in FDA notices) is that certain error codes (identified by rule numbers) trigger automatic technical rejection, requiring the sponsor to fix and resubmit (www.fda.gov) (en.ennov.com). For example, an FDA guidance notice (Federal Register, Aug 2021) announced that eCTD validation errors 1551 and 1553 would be enforced as high severity, so a submission failing them is rejected unless corrected (www.fda.gov). The exact enforcement mechanism varies by agency: some may allow minor issues to be resolved by a brief inquiry, but on core technical failures there is no workaround besides resubmission.

If validation passes or has only “low” issues, the submission is technically accepted and then routed to reviewers. Even then, some “medium” issues (rating semantics or formatting nuances) might not block filing but could later prompt queries from reviewers. High-severity validation issues, however, must be corrected beforehand.

In practice, the **validation criteria** each agency uses are often **publicly documented** or derived from published specifications. The ICH itself has published technical documents (e.g. the eCTD backbone DTD and stylesheets) and a Q&A wiki on the ICH.org site. Agencies typically maintain a “submission standards” repository: for example, FDA’s eCTD Submission Standards page lists the *Specifications for eCTD Validation Criteria* (version 4.6 as listed in August 2026) and an eCTD Conformance Guide (www.fda.gov) (www.fda.gov). Health Canada publishes its validation rule set in a guidance document (the excerpt in the last section is from the **Health Canada eCTD Validation Rules v5.3**).

Each validation rule has an identifier (like A01, B17, etc. in Health Canada’s schema) and a defined severity (Error or Warning). Errors must be fixed; warnings should be reviewed. For example, Health Canada rule **A03b** flags any PDF >200MB as an Error, while **A03a** only warns for 150–200MB (www.canada.ca). Table 1 (below) illustrates common FDA-grade errors and their causes (drawn from FDA training materials)，and Table 2 shows sample rules from a Health Canada v5.3 eCTD rule set. These demonstrate how precisely errors are defined and categorized.

Error / Symptom	Possible Cause (Illustrative)
Duplicate sequence number	The folder sequence (e.g. 0002) has already been submitted previously (www.accessdata.fda.gov).
Missing backbone XML	No `index.xml` (or regional XML) was included, or it had a bad name/path. (www.accessdata.fda.gov)
No files (“No data”)	The submission folder contained no document files; essentially empty content (www.accessdata.fda.gov).
Corrupted media	Files could not be read (broken PDFs, incomplete zip, etc.) (www.accessdata.fda.gov).
Form mismatch	The eCTD’s application number in XML didn’t match the number on the FDA form (www.accessdata.fda.gov).
Sent to wrong Center	File was routed to CDER instead of to CBER (or vice versa), based on form type (www.accessdata.fda.gov).

Table 1: Example FDA eCTD rejection errors and root causes (www.accessdata.fda.gov). A single such high-level error (marked as High severity) causes FDA to consider the submission not received.

Rule ID	Rule Name	Severity	Description (from Health Canada eCTD v5.3)
A01	Empty Folders	Error	Flags any folders in the sequence with no content (www.canada.ca).
A03a	File Size (warning)	Warning	Warns if PDF between 150–200MB (or other non-XPT files over 100MB) (www.canada.ca).
A03b	File Size (error)	Error	Errors if PDF >200MB or SAS XPT >1GB (www.canada.ca).

Rule ID	Rule Name	Severity	Description (from Health Canada eCTD v5.3)
A05a	Sequence Requirements	Error	Initial sequence must be "0000"; any misnumbering flags error (www.canada.ca).
A05b	Higher sequences found	Error	Error if the validated sequence is not the highest sequence in the application folder (gaps in numbering are caught by rule A07).
A06a	XML Backbone Identification	Error	Checks for the presence and validity of the XML index (ICH + regional) (www.canada.ca).
A06b	STF Identification	Error	Checks presence and naming of Study Tag Files (STFXML) if provided (www.canada.ca).

Table 2: Sample eCTD validation rules from Health Canada's v5.3 rule set (www.canada.ca) (www.canada.ca). Fields include rule ID, an error/warning flag, and description. (Health Canada uses letter prefixes to group rules; A–E are general checks, F–I file-level checks, etc.)

These tables illustrate the granularity of validation: missing a single document, using an unsupported file extension, or even just misnaming a file can trigger an immediate error. Sponsors therefore must rigorously use validation tools. As an industry whitepaper notes, *"the simplest deviation … can make your eCTD submission invalid, which would mean timely and costly resubmission."* (www.onixls.com) Typical pitfalls include **incorrect PDF formatting** (embedding fonts, fonts password-protected, or wrong version) and **navigation errors** (broken hyperlinks/bookmarks in PDFs) (www.onixls.com).

Validation Tools and Industry Practice

Because manual checking of thousands of files is impractical, sponsors rely on specialized software solutions from life sciences vendors. Some popular eCTD publishing and QC products include the Lorenz DocuBridge/eValidator suite (used by FDA), EXTEDO eCTDmanager, ArisGlobal eCTD Solutions, PAREXEL eCTD tools, and OpenText/LibreHealth platforms. These tools incorporate the official validation logic and often allow the user to select which regional specification to validate against (US, EU, Japan, etc.).

For example, FDA has published that it uses **Lorenz eValidator** (version 25.2 as listed in August 2026) for eCTD submissions (both v3.2.2 and v4.0) (www.fda.gov) (www.fda.gov). Pinnacle 21 Enterprise is also listed on FDA's site, but it is used for clinical data (SDTM, SEND) validation, not for eCTD structure (www.fda.gov). Health Canada suggests sponsors "use a commercially available tool to validate their regulatory transactions in eCTD format, prior to filing them to Health Canada" (www.canada.ca). Indeed, many tools allow batch validation of entire application branches or cross-sequence linking to catch issues early.

Even small companies often outsource their final validation. The cost of a late-stage validation error is high: aside from submission fees lost, any delay in drug approval timeline is extremely costly (www.docshifter.com). Similarly, regulators provide compliance help; the FDA's Small Business Assistance center has web-based training indicating exactly which structures cause RTF (www.accessdata.fda.gov), and the EMA holds periodic webinars on eCTD tool usage and common errors.

Table 1 above compiles a few of the **most common error outputs** by FDA's eCTD validation tool. Consulting firms warn that **formatting consistency** is critical: eCTD documents "must adhere to specified formats" or they will be flagged (www.onixls.com). Commonly cited failures include using wrong PDF versions, having incorrect bookmarks, or omitting required appendices. A consulting blog reports that *"recurrent eCTD validation criteria changes mean eCTD validation issues are particularly common"* (www.onixls.com). These sources stress that staying up-to-date on the latest criteria is a continuous task – a missed update can invalidate an entire package.

Overall, eCTD validation is not just a formality but an integral part of submission quality assurance. Proper validation prevents wasted review cycles and ensures the agency receives a consistent, review-ready dossier.

Regulatory Perspectives and Requirements

Each regulatory agency has codified its eCTD expectations in guidance documents or technical specifications. While based on the ICH eCTD spec, regional differences (mostly Module 1 requirements) mean that validation must be done against **region-specific criteria**. Below we outline leading authorities' approaches:

- **United States (FDA)**: CDER and CBER have progressively tightened requirements. Earlier PDUFA-era plans anticipated phased eCTD timelines (www.accessdata.fda.gov), but eCTD became mandatory only from May 5, 2017 for NDAs, ANDAs and BLAs and from May 5, 2018 for commercial INDs and master files (www.fda.gov). The FDA publishes a comprehensive *"eCTD Technical Conformance Guide"* and *"Submission Standards"* web pages (as seen in the FDA's eCTD v3.2.2 and v4.0 standards tables (www.fda.gov) (www.fda.gov)). These list every relevant document: eCTD backbone specs (ICH M2), US Module 1 schema, stylesheets, controlled vocabularies, etc. The FDA also regularly issues Federal Register notices updating policies (for example, announcing support of eCTD v4.0 for new applications from Sep 16, 2024 (www.fda.gov) (www.fda.gov), and specifying which validation rule codes are high-severity (www.fda.gov)).

The FDA's enforcement is strict.A small business guidance even explicitly lists results of failed checks (as in Table 1) to warn filers (www.accessdata.fda.gov). The bottom line: FDA expects flawless technical compliance, supported by their own validator.

- **European Union (EMA and National Competent Authorities)**: EMA and EU member states follow ICH eCTD v3.2.2 (Module 2–5) plus an EU-specific Module 1 outline. Historically, Europe also accepted a transitional *NeeS* (non-eCTD electronic) format until around 2015; post-2010, however, centralized (CAP) filings must be eCTD. EMA's *eSubmission* website publishes the official **EU Module 1 Specification** and **Validation Criteria** (organized by version numbers). For instance, as of mid-2025 EMA was implementing *Module 1 v3.1.1* and its associated Validation Criteria *v8.2* (esubmission.ema.europa.eu). The EMA site explicitly announces effective dates: Module 1 v3.1 was accepted Oct 2024, with mandatory submission compliance by Mar 2025, and the v3.1.1/8.2 package accepted Oct 2025 with mandatory use by Dec 2025 (esubmission.ema.europa.eu) (esubmission.ema.europa.eu). Such communications highlight that *both* the Module 1 structure and its validation rules evolve over time.

EMA's Harmonisation Group also contributes guidance (e.g. the *Harmonised Guidance eCTD* documents) that influence validation. National authorities (e.g. MHRA in the UK, Swissmedic, etc.) may have slight variations, but generally follow the EU M1 spec and rules. EMA validations similarly produce detailed reports listing missing or misplaced documents. Anecdotally, European agencies will refuse to accept an MAA if key pieces (e.g. EU Product Information, or missing eAF, or non-compliant cover letter) are absent, analogous to FDA.

- **Canada (Health Canada)**: Health Canada requires either eCTD or an older non-eCTD electronic format (the non-eCTD electronic-only format) depending on product category (www.canada.ca). It provides detailed *Guidance on Preparation of Drug Regulatory Activities in eCTD Format* and implements eCTD v3.2.2 (with Canadian Module 1 v2.0) for drugs. Health Canada's standalone "Validation Rules" guidance version 5.3 took effect on May 31, 2025 (www.canada.ca). This explicitly aligns with ICH v3.2.2 and the Health Canada Module 1 guidance. It emphasizes the goal: "to help ensure Sponsors provide a valid electronic transaction to Health Canada, and reduce errors and follow-up" (www.canada.ca). As shown in Table 2, the HC rules enumerate dozens of specific checks. Health Canada states that every transaction is validated on receipt; if validation fails, they will send a comprehensive report of each error to the sponsor (www.canada.ca). This is a firm policy of feedback. In case of failure, the sponsor must fix and refile (there is no remediation of an active file).
- **Japan (PMDA)**: Japan's Pharmaceuticals and Medical Devices Agency has also moved to eCTD compliance. PMDA accepted eCTD v3.2.2 for approval applications made through March 31, 2026, after which applications move to eCTD v4.0 (www.pmda.go.jp). Local guidance documents outline Module 1J (Japan-specific) and require Japanese translations of key modules. Validation rules for PMDA are not as readily published as FDA's or EMA's, but sponsoring companies invariably use the same global publishing tools to satisfy PMDA eCTD checks.
- **Other Authorities (WHO, ICH regions)**: Many countries follow either ICH standards or mirror EMA/Health Canada style guidelines. For example, the WHO, and regions like APAC (Australia, Singapore), and the GCC have their own timelines for adopting eCTD. Each will have its own formal or informal checklist (many adopt ICH CTOC headings, for instance) and some publish validation rules. Switzerland (Swissmedic) runs its own eCTD process with a Swiss Module 1, closely aligned with the EU approach.

In **summary**, regulators require strict technical conformance to eCTD specifications. They periodically update rules, which sponsors must follow. The common thread is that one must validate against *the exact criteria in effect at submission time* using tools reflecting those criteria. Failure to do so typically leads to rejection.

Data and Trends in eCTD Usage and Validation

The transition to eCTD has resulted in an enormous volume of electronic submissions. For context, the FDA's electronic gateway (ESG) statistics show that from 2014–2025 **approximately 66 million** submissions were exchanged across all centers (CDER, CBER, CDRH, etc.) (www.fda.gov). (This includes everything from small

data sets to full NDAs, but the majority are likely routine communications and safety reports). CDER alone processed ~2.84 million electronic transactions in that period (www.fda.gov). These totals illustrate the scale of electronic exchange, but they are not a denominator for eCTD rejections.

A key metric of validation is the **rejection rate**. According to a 2023 conference presentation (cited in industry press), FDA's rejection rate for submissions in eCTD format was reported as *"less than 2%"* (en.ennov.com). This figure is low, and because FDA's gateway totals include many non-eCTD transaction types, it cannot be converted into an absolute count of eCTD rejections. This is consistent with anecdotal industry experience that most companies catch issues pre-submission through QA, but a few errors do slip through. The broken-out reasons (see Table 1) show that many rejections are ultimately preventable technical glitches.

Industry surveys also reveal some trends in handling validation. Many companies maintain dedicated publishing/validation teams or engage specialized vendors. Larger corporations often have in-house software and processes to pre-validate. The job of ensuring a "clean" eCTD is sometimes described as requiring dozens of person-hours per submission. Consulting blogs note **common pitfalls**: besides format errors, *incomplete submissions* (e.g. missing required documents or adding them in the wrong sequence) can cause the dossier to be marked "incomplete" by the agency's system (www.onixls.com).

Quantitative studies on this topic are scarce (likely proprietary), but some data exist. Delays are costly, as noted above (www.docshifter.com). Given vendor estimates that a day of delay can cost up to $8M (www.docshifter.com), companies have strong financial incentive to minimize validation failures. That economic pressure ensures that most sponsors invest in thorough validation - supported by the fact that FDA refunds only 75% of the application fee when it refuses to file an application, so the sponsor forfeits 25% (about $1.15 million of the FY2027 fee of $4,600,753 for an application requiring clinical data (www.fda.gov)) (www.docshifter.com).

From a regulatory perspective, rigorous validation has delivered efficiencies. Agencies no longer need to manually inventory submissions for completeness – they rely on the automated engine. This has presumably reduced initial triage times compared to the paper era (though data quantifying this are not publicly available). The volume of eCTD filings has continued to grow (with every new IND/MAA, as well as supplements and even NDAs for minor changes, being filed electronically). This trend is only accelerating: for example, FDA's table shows a steady increase in annual submissions up to 2024 (www.fda.gov).

Finally, statistical analysis of *types of errors* shows the most common culprits. In the FDA slide deck (Ennov blog), the top technical errors were often duplicate sequence, missing backbone file, wrong Form usage, etc. Collectively, the data underscore that *process discipline* in assembling submissions is key, and that the cost of mistakes is substantial.

Case Studies and Examples

To illustrate the real-world impact of validation requirements, consider a few scenarios. The EMA item describes a published EMA notice; the others are hypothetical composites built on published rules, not documented sponsor events:

- **FDA – Duplicate Sequence (hypothetical)**: A small biotech plans to submit Amendment 2 of an IND (as sequence 0002). However, due to a coordinator oversight, they numbered the folder `0001` again (which had already been used). FDA's validator immediately flagged "Duplicate sequence submission" since 0001 was previously received. The error report (analogous to Table 1) was sent back, and the submission was considered not filed (www.accessdata.fda.gov). The company reverted to draft, correctly numbered the folder as 0002, and resubmitted. This common mistake underscores why sequence numbering is strictly enforced (Health Canada rules A05a/A07 likewise prohibit gaps or repeats (www.canada.ca)).
- **EU (EMA) – Missing Tracking Table**: EMA uses a tracking table to index certain critical documents in Module 1 (e.g., lists of pediatric studies). In 2025, after updating Validation Criteria to version 8.1, EMA found that some submissions failed two new rules (15.11 and 15.12) regarding a mandatory tracking table (esubmission.ema.europa.eu). However, these rules were not meant to apply to EDQM (pharmacopoeial) submissions. The EMA temporarily allowed EDQM filings to pass even if those rules were violated (esubmission.ema.europa.eu). This example shows how subtle policy nuances can be embedded in validation rules, and how agencies may adjust criteria in response to stakeholder feedback.
- **Health Canada – File Size Limits (hypothetical)**: A Canadian generic drug sponsor prepares an eCTD submission with many large analytical reports. Two files slightly exceeded 200 MB after conversion, triggering Health Canada's A03b rule (www.canada.ca). Although the original PDFs rendered fine, the validator flagged them as errors (no workaround; PDf had to be split into two volumes). The sponsor had to rebuild these documents split by section, then republish the whole application. This illustrates how technical constraints (200MB limit) can force extra work if not anticipated.
- **Japan – Module 1 Error (hypothetical)**: A firm submitting to Japan's PMDA reuses a Module 1 from a US NDA with translations, but fails to update the regional metadata per Japan's requirements. The submission fails PMDA's checks for incorrect regional metadata, an error of metadata rather than document content. The company had to revise the Module 1 XML to meet Japan's schema. This type of region-specific validation (PMDA has its own attribute lists) highlights that multi-country submissions must run all relevant regional checks before filing.
- **Cross-Agency Tool Implementation (hypothetical)**: A sponsor develops an internal eCTD publishing system. During user acceptance testing, the team reran an older FDA-approved NDA through the new tool chain – only to find 50+ low-severity warnings due to slight XML formatting differences. Although none were fatal, their system flagged each as potential problems. The team then reviewed FDA's *Technical Conformance Guide* to adjust their XML generator. This case shows that even "low-level" rules (like whitespace or attribute ordering) matter in some validators.

In each of these examples, the *validation process* functioned as intended: catching deviations before regulators had to manually identify them. Had such checks not been automated, reviewers might have been confused by an otherwise valid submission or, worse, missed key errors. Conversely, the scenarios also illustrate the frustration and cost for sponsors: every error, even trivial, requires a cycle of correction and resubmission, delaying review.

Some lessons from these experiences:

- **Pre-submission validation is indispensable.** Experienced publishing teams typically run full validation as soon as their document compilation is “mostly ready,” and again after final builds. Early catching of missing files is much cheaper than a late RTF.
- **Keep close watch on rule updates.** Several cases arose simply because the criteria version had changed since last submission (e.g. when FDA updated the accepted PDF versions or when EMA switched Module 1 versions). Sponsors often assign an “eCTD quality lead” to monitor agency announcements (see EMA’s news feed and FDA notices).
- **Invest in staff training.** Teams that deeply understand the structure (and have checklists) can avoid many pitfalls. eCTD publishing and validation is widely treated as a specialized skill set.
- **Use multiple perspectives.** Some companies integrate validators from different vendors (though all should follow the same rule set) to catch any discrepancies. Others bring in mock regulatory reviewers to spot anything the machine might not catch (for example, whether section titles make sense).

These case studies and practices underscore that eCTD validation is both a technical challenge and part of a sponsor’s regulatory strategy. Errors caught early can be fixed quietly; errors caught at FDA/EMA front desks cost time and money. Next we discuss the broader interpretation and implications of these requirements.

Discussion: Implications and Future Directions

The comprehensive validation requirements for eCTD carry significant implications for the pharmaceutical industry, regulators, and the future of drug development.

Quality and Efficiency Trade-offs

On the positive side, stringent validation ensures a baseline quality. By forcing compliance at the technical level, regulators greatly reduce the chance that reviewers waste time on incomplete or malformed dossiers. For instance, an eCTD missing the entire Nonclinical Report is a much bigger mistake if only discovered during review, rather than at intake. Automated checks remove many trivial questions (e.g. “Where’s the signature page?”) and let reviewers focus on substantive content. In aggregate, industry and agency surveys suggest that standardized electronic submissions speed up the filing process by reducing back-and-forth, although quantifying this precisely is hard.

However, there is a cost. **Validation complexity** has grown with each new version of the spec. Sponsors must continuously invest resources (people, systems) to track updates across FDA, EMA, Health Canada, Japan, etc., and to update their submission processes. Smaller companies without dedicated regulatory publishing groups may rely on consultants or consultants, which adds expense. Even large companies build automated pipelines to ensure consistency because manual compilation is prone to error.

The differing U.S. and EU Module 1 requirements illustrate a broader theme: *divergent rules* still exist, making global submissions complex. Although both FDA and EMA use eCTD, the exact metadata atoms differ (e.g., `<promotional-material-type>` is in FDA but not in EMA's M1). These differences mean validation cannot be entirely globalized; many firms must produce separate eCTDs for the EU and US. There is ongoing work under ICH M8 (eCTD v4) to harmonize more of the vocabularies, but complete alignment remains challenging.

Sponsors weigh the benefit of catching errors early against the burden of over-validation. For instance, practitioners note that receiving dozens of benign warnings (on image quality or file metadata) can clutter the report and obscure real problems. There is an argument (among some in industry) that extremely strict enforcement may not add safety value in all areas—such as limiting file sizes to 200MB (an arbitrary technical constraint). On the other hand, regulators view these constraints as necessary for managing their submission review systems at scale and over long archiving periods. Thus, one tension is between *practicality* (flexibility, sponsor convenience) and *rigor* (unambiguous technical compliance). At present, regulators clearly lean toward maximum rigor.

The consequences of failure reflect this. As a DocShifter analysis notes, even a minor RTF can cost millions in lost fees and market opportunities (www.docshifter.com). This high-stakes environment means the industry often accepts strict rules as part of the regulatory burden. Nonetheless, there have been industry calls for improved guidance or grace periods when new rule versions roll out. For example, EMA's phased acceptance of multiple spec versions in transition periods (as seen in June–Sep 2025) (esubmission.ema.europa.eu) was one response to allow sponsors time to adapt. Such policies are important to mitigate risk for filings in progress.

Trends and Future Outlook

Moving to eCTD v4.0: The next big change is the global migration to *eCTD version 4*. ICH M8 defines the v4 specification, and FDA began accepting v4 for new applications in Sep 2024 (www.fda.gov). EMA allows optional v4 use for new centrally authorised product MAAs from Dec 22, 2025, with prior contact with its eCTD v4.0 team (esubmission.ema.europa.eu) (esubmission.ema.europa.eu), and other agencies are moving too (e.g. Japan, where PMDA accepted v3.2.2 for approval applications only through March 31, 2026 (www.pmda.go.jp)). As of mid-2026, FDA has not yet announced a date by which eCTD v4.0 will become mandatory for new NDAs/BLAs/INDs — v3.2.2 remains the dominant production format, with v4.0 uptake still in early adopter phase. eCTD 4 adds some new features: a modernized XML message based on HL7 Regulated Product Submissions (RPS), expanded use of controlled vocabulary, and better support for documents (like modules for device combination products, or dynamic documents). Validation for v4 includes all existing checks plus new ones on the additional XML fields (e.g. new Module 3 data points). Sponsors must adapt their publishing tools and QC processes accordingly. Notably, FDA has already published *"Specifications for eCTD v4.0 Validation Criteria"* (ready by Oct 2025) (www.fda.gov). The Lorenz validator is updated for v4 as well.

Integration with Regulatory IT Systems: eCTD is one pillar of the agencies' IT modernization. Others include regulatory information management (RIM) systems that track applications, electronic data standards (like CDISC), and portals that allow interactive queries. There is talk of going beyond eCTD for some raw data types: for example, FDA now requires standardized datasets (SDSs) alongside eCTD for clinical trials (like

SDTM/XPT packages) and has separate validators for that (Pinnacle21). In the future, portions of submissions might shift to database entries or JSON payloads (especially with HL7/FHIR initiatives in regulatory reporting). However, any such changes will still need validation analogous to eCTD.

Automation and AI: Given the repetitive nature of validation, sponsors are exploring further automation. Some emerging tools can semi-automatically fix eCTD errors (e.g. renaming files, re-generating broken indexes) or check consistency across documents using text analysis. There is interest in using AI to predict possible missing content (by semantic analysis of dossier text) or to cross-validate data in Module 2 summaries vs Module 5 reports. While still early, such technologies could reduce the manual effort in pre-submission QC. Agencies might also employ machine learning to review narrative content, but for validation specifically, rule-based checks dominate due to the deterministic nature of the requirements.

Global Harmonization: Longer-term, regulators aim for even more convergence. ICH working groups continue to refine guidelines so that, for example, the Non-eCTD electronic Submission (NeeS) format is fully phased out and only eCTD remains. Even within eCTD, efforts to harmonize Module 1 (through M8 implementation guides) will simplify multinational filings. For now, the industry is focused on digesting eCTD v4.

Impact of COVID-19 and Remote Work: The pandemic accelerated digital collaboration. Many companies had to finalize eCTD submissions remotely, relying entirely on electronic workflows. This broadly validated the utility of eCTD – submissions could proceed uninterrupted when offices closed. It also highlighted the importance of robust electronic quality systems. Regulators similarly moved more quickly to accept e-submissions for COVID therapeutics, sometimes issuing temporary guidances. These developments strengthen the expectation that future regulatory communications will remain heavily digital, further cementing eCTD as the norm.

Conclusion

The eCTD format has transformed pharmaceutical regulation by making submissions digital, standardized, and shareable. Its **validation requirements** are correspondingly strict, reflecting the need for completeness, accuracy, and interoperability in electronic dossiers. This report has outlined the technical and regulatory landscape: how eCTD is structured, the specific rules imposed by major agencies, the tools used by sponsors to comply, and the data on how often submissions fail these checks.

Key points include:

- The **historical evolution** from CTD to eCTD laid the groundwork for today's systems. With CTD harmonized globally, the move to eCTD leveraged common data structures but imposed new technical constraints (XML indexes, approved file formats).
- **Regulatory rules** are detailed and continuously updated. Agencies like FDA, EMA, and Health Canada each publish their own validation criteria in harmony with ICH documents. These rules cover everything from folder names to file sizes to metadata attributes, as documented in Sections 3–5.

- Sponsors bear the onus of compliance. They must use updated validation tools (Lorenz, etc.) to pass the same checks, because high-severity errors can lead to outright rejection. Industry guidance universally emphasizes early and repeated validation (www.onixls.com) (en.ennov.com).
- **Impact on industry** is non-trivial. Even with generally low rejection rates (~2% as per FDA) (en.ennov.com), each error costs time and money. Refusals may forfeit part of the application fee and delay market entry, which one vendor estimates can cost millions of dollars per day (www.docshifter.com).
- **Trends**: eCTD is expanding (v4 rollout) and being integrated into broader digital strategies. Regulators and companies alike are investing in automation to make validation more efficient. Global harmonization efforts continue but differences persist in operational details.

Looking forward, the eCTD validation landscape will remain dynamic. The introduction of eCTD 4.0 necessitates another cycle of learning and system upgrades. Agencies are likely to refine validation criteria to keep pace with new submission modalities. On the horizon, linking eCTDs with real-time data (e.g. continuous monitoring submissions, or integrating patient registries) will require new kinds of checks. However, the core principle will survive: **"Only well-formed, standards-compliant eCTD submissions will be accepted."** This ensures regulatory review resources focus on science and safety, rather than on deciphering flawed formats.

The exhaustive requirements documented here – across hundreds of rules and multiple regions – underscore a key conclusion: eCTD validation is not a peripheral concern but a central pillar of the regulatory process. Sponsors and regulators must both maintain vigilance to its evolving demands. By investing in robust validation practices, companies protect their submission investments and facilitate faster, more reliable regulatory decisions – ultimately advancing patient access to therapies.

Sources: Authoritative regulatory guidance and industry literature as cited throughout, including FDA eCTD specifications (www.fda.gov) (www.fda.gov), EMA eSubmission updates (esubmission.ema.europa.eu) (esubmission.ema.europa.eu), Health Canada rules (www.canada.ca) (www.canada.ca), and expert commentary (pharmafile.com) (www.onixls.com) (en.ennov.com) (www.docshifter.com). These sources underpin all claims in this report.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://en.ennov.com/blog/reasons-ectd-rejection-fda/
2. https://www.fda.gov/industry/resources/submission-statistics
3. https://www.docshifter.com/blog/fda-refuse-to-file/
4. https://www.canada.ca/en/health-canada/services/drugs-health-products/drug-products/applications-submissions/guidance-documents/ectd/notice-validation-rules-regulatory-transactions-submitted-health-canada-electronic-common-technical-document-format-2016-12-1.html
5. https://www.accessdata.fda.gov/scripts/cder/training/eCTD/validate/validate/fd_06_01_0040.htm
6. https://www.fda.gov/drugs/electronic-regulatory-submission-and-review/ectd-important-notices
7. https://www.fda.gov/drugs/electronic-regulatory-submission-and-review/ectd-submission-standards-ectd-v40-and-regional-m1
8. https://esubmission.ema.europa.eu/eCTD%20NMV/eCTD.html

9. https://esubmission.ema.europa.eu/eumodule1/index.htm
10. https://pharmafile.com/features/evolution-ectd/
11. https://www.ema.europa.eu/en/ich-m2-electronic-common-technical-document-ectd-scientific-guideline
12. https://studylib.net/doc/25507272/ectd-tech-guide-v1.5-final
13. https://www.fda.gov/drugs/electronic-regulatory-submission-and-review/ectd-submission-standards-ectd-v322-and-regional-m1
14. https://www.accessdata.fda.gov/scripts/cder/training/ectd/submit/submit/fd_03_01_0020.htm
15. https://www.canada.ca/en/health-canada/services/drugs-health-products/drug-products/applications-submissions/guidance-documents/common-technical-document/updated-guidance-document-preparation-regulatory-activities-non-ectd-electronic-only-format.html
16. https://www.fda.gov/media/76797/download
17. https://www.onixls.com/2023/01/ectd-validation-tips/
18. https://www.pmda.go.jp/int-activities/int-harmony/ich/0120.html
19. https://www.fda.gov/prescription-drug-user-fee-act-pdufa

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore AI enablement, pharma and biotech software development, data engineering and BI, and Veeva services. Contact IntuitionLabs to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: https://intuitionlabs.ai

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.